DOI:
10.1039/c7ob02709e

Document Version
Peer reviewed version

Journal Name

COMMUNICATION

Methylene-linked *bis*-phenylbenzimidazoles as a new scaffold for interaction with the telomeric DNA/RNA hybrid duplex

M. K. Islam,[a,b] P. J. M. Jackson,[a,c] D. E. Thurston[a]* and K. M. Rahman[a]*

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org

We report here a series of novel methylene-linked *bis*-phenylbenzimidazoles intercalators that stabilize telomeric DNA/RNA hybrid (tDRH) structures by up to 7.2°C at a 1 µM ligand concentration while having negligible affinity for DNA/DNA duplexes, although with a low affinity for quadruplex DNA. We have used molecular modelling studies to rationalize this selectivity, concluding that the methylene spacer between the terminal benzimidazole and phenylene moieties plays a key role in facilitating the *bis*-intercalating process.

The DNA/RNA hybrid (DRH) duplex structure (**Figure 1b**) was first proposed six years after the double-helical structure of DNA (**Figure 1a**) was reported by Watson and Crick.[1] The first DRH duplex was identified in 1961 by annealing a RNA strand with a complementary DNA strand[2], and the first DRH duplex was synthesized in 1960 by reacting oligodeoxythymidylic acid with polyriboadenylic acid[3]. In 1967, X-ray diffraction and CD spectroscopy studies confirmed the different conformation of the RNA/DNA hybrid structure compared to duplex DNA.[4]

In cells, DRH duplex formation is an important component of the mechanism for elongation of the telomeric DNA sequence at the ends of chromosomes. Telomeres exist as a protein-DNA composite with long repeats of a unique six base sequence (5´-TTAGGG). Upon activation of telomerase, the protein hTERT, which contains an RNA sequence (3´-CAAUCCCAAUC-5´) as part of its structure, forms an RNA-DNA duplex with one of the repeating telomeric 5´-TTAGGG sequences on the 3´-strand of a chromosome.[5] DNA polymerization then proceeds *via* reverse transcription of part of the RNA template to synthesize one telomeric repeat (TTAGGG) on the 3´-end of the DNA primer. The DRH duplex formed is thus a unique structure in cells, and is considered a high-value drug target in oncology.[6] One therapeutic approach is to identify small molecules capable of stabilizing the tDRH duplex, thus preventing telomere extension. Other approaches have attempted to modulate substrate/enzyme interaction, and/or inhibit dissociation of the enzyme from the substrate.[7, 8]

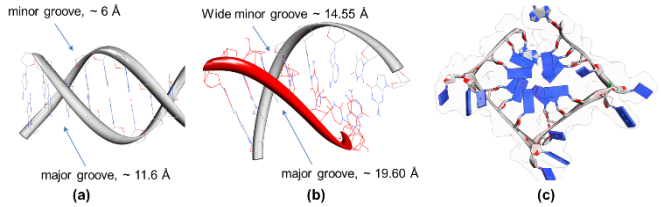

Figure 1: Common forms of DNA: **(a)** DNA/DNA Duplex (B-Form), **(b)** DNA/RNA hybrid duplex (A-Form), and **(c)** DNA quadruplex.

A number of molecules have been reported to bind to non-telomeric DRH duplexes such as ethidium derivatives, ellipticine, paramomycin, ribostamycin and neomycin (**S1, ESI**).[9-13] In 2010, Wheelhouse and co-workers reported a pyrimidine-connected *bis*-sulfane molecule (**Figure 2**) with a 20-fold preference for binding to a poly(dA)-poly(rU) hybrid duplex compared to an equivalent RNA fragment, a 3-fold preference over duplex DNA, and 7-fold preference over the alternative poly(rA)-poly(dT) hybrid sequence, in a competition dialysis assay.[13]

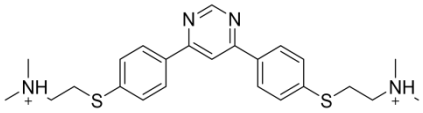

Figure 2: Molecule reported by Wheelhouse and co-workers with an affinity for poly(dA)-poly(rU) hybrid duplex.[13]

Through screening 2307 molecules from the NCI's Diversity Set II, Natural Products Set II and Mechanistic Diversity Set libraries against a telomeric DRH duplex sequence (tDRH, *i.e.*, 5´-TTA-GGG-TTA-GGG-TTT-TTT-CCC-UAA-CCC-UAA-

[a]*Institute of Pharmaceutical Science, School of Cancer and Pharmaceutical Sciences, King's College London, London SE1 9NH, United Kingdom.*
[b]*Department of Pharmaceutical Chemistry, University of Dhaka, Dhaka-1000, Bangladesh.*
[c]*Presently at Femtogenix Ltd, BioPark, Broadwater Road, Welwyn Garden City AL7 3AX, United Kingdom.*
Email: k.miraz.rahman@kcl.ac.uk (KMR), david.thurston@kcl.ac.uk (DET); Tel: +44 (0) 207 848 7507
♦Electronic Supplementary Information (ESI) available: Experimental procedures, and supplemental tables, spectra and graphs. See DOI: 10.1039/x0xx00000x

3´), we identified a *bis*-2-methylquinoline structure (NSC273829) (**Figure 3**) capable of stabilizing the telomeric DRH sequence by 11.5°C and 3°C at 5 μM and 1 μM, respectively, but with negligible affinity for DNA duplex (*i.e.*, ΔT_m = 0.6°C for a 1 μM ligand concentration). This compound was one of a set of thirteen molecules previously reported to have G-quadruplex binding affinity.[14] The compound was originally reported by the NCI to have anticancer activity in mouse L1210 Leukaemia xenografts[15] and to be active against leishmania[16].

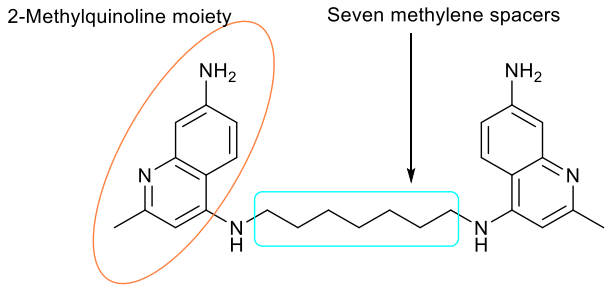

Figure 3: Structural features of NSC273829.

Our molecular modelling studies suggested that NSC273829 has the potential to interact with the tDRH duplex through its minor groove, but it does not have the appropriate 3-dimensional shape for intercalation unlike previously reported tDRH targeting molecules such as the ethidium derivatives and the ellipticines. Therefore, modelling was used to design analogues of NSC273829 potentially capable of a greater affinity with the tDRH duplex through an intercalative mechanism. This was achieved by incorporating a methylene-based spacer between the terminal *bis*-intercalating groups that included phenyl groups to enhance intercalation and to facilitate orientation of the terminal benzimidazoles toward the DNA bases. In an initial study, molecular modelling suggested that the 1-benzyl-1*H*-benzo[*d*]imidazole moiety could fit into the telomeric DNA/RNA hybrid duplex. A benzimidazole moiety was also chosen because this class of heterocycle is associated with a wide spectrum of biological activities.[17-19] Moreover, benzimidazoles offer additional hydrogen binding opportunities with DNA/RNA bases compared to the quinoline moiety present in NSC273829. A set of molecules (**Library 1**) was designed using this approach, with the 1-benzyl-1*H*-benzo[*d*]imidazole units linked by methylene spacers containing 5-10 carbons (**Figure 4**).

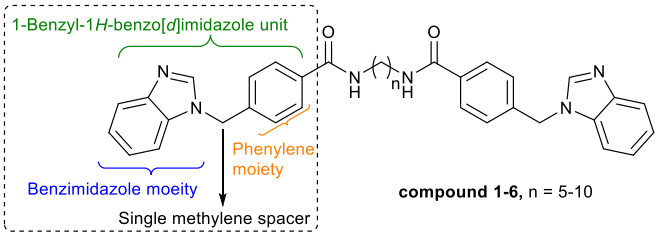

Figure 4: Structural features of Library 1 molecules.

All Library 1 molecules were synthesized through a simple amide coupling reaction. A solution of 4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzoic acid (0.20 mmol) in dimethylformamide (7 mL) was treated with 1-hydroxybenzotriazole (0.40 mmol) and N,N'-diisopropylcarbodiimide (0.35 mmol) at room temperature. After an initial activation step to form the activated acid ester (usually 20-30 minutes), the respective diamines containing 4 to 10 methylene spacers (0.24 mmol) were added to the reaction mixture, which was allowed to stir overnight (14-15 h). Upon confirmation of product formation by LC-MS, the reaction was quenched with water (10 mL) and extracted with ethyl acetate (3 x 20 mL) (**S2**, **ESI**). All final products were purified by column chromatography, and characterised by NMR and MS, and were >95% pure based on analysis in two solvent systems using LC-MS.

A FRET melting assay (**S3, ESI**) was carried out as a preliminary biophysical screen, and compound **4** (containing eight methylenes; **Figure 4**, n=8) was observed to have a 9-fold selectivity for the tDRH duplex compared to the control DNA duplex (cDD), stabilizing the hybrid duplex by ΔT_m values of 9.5 and 7.2°C at 2 μM and 1 μM, respectively (**Table 1**). It was observed from these studies that molecules containing an even number of methylene groups linking the benzimidazole moieties stabilised the tDRH to a greater extent. For example, in the case of compound **4** (eight methylenes), it provided greater stabilization than **5 (**nine methylenes), with a difference of 2.7°C observed (*i.e.*, 7.2°C compared to 4.5°C). A similar pattern was observed for compound **6** (containing ten methylenes) compared to **5**, with **6** stabilising tDRH to a greater extent (*i.e.*, 6.8°C for **6** compared to 4.5°C for **5**, a difference of 1.7°C). Compounds were also screened for binding to the RNA/RNA duplex (RRH; *i.e.*, 5´-UUA-GGG-UUA-GGG-UUU-UUU-CCC-UAA-CCC-UAA-3´), but failed to effect significant stabilisation (*i.e.*, ΔT_m values ≤0.5°C).

Table 1: Comparison of analogues in terms of FRET and CD studies.

Library	Compound	tDRH		F21T		cDD	
		ΔT_m	CD	ΔT_m	CD	ΔT_m	CD
NCI	NSC273829	3.2	-	11.8	-	0.6	-
1	**1**	2.4	--	5.8	--	0.6	--
1	**3**	4.5	--	5.5	--	1.3	--
1	**4**	7.2	++	6.5	--	0.8	--
1	**5**	4.5	--	6.5	--	1.2	--
1	**6**	6.8	++	6.2	--	0.9	--
2	**7**	0.5	--	0.2	--	0.3	--
3	**12**	0.4	--	0.3	--	0.2	--
4	**17**	0.5	--	0.5	--	0.3	--

ΔT_m are at 1 μM ligand concentration, '++' indicates CD shifts, '--' indicates no shifts in the CD analysis, and "-" indicates that the compound was not evaluated. Data for 2 μM in S3, ESI.

Library 1 molecules were also screened in the FRET assay for binding to the telomeric quadruplex sequence (F21T, 5´d-FAM-GGG-TTA-GGG-TTA-GGG-TTA-GGG-TAMRA-3´) (**Table 1**). The most active molecule (**4**) doubled the stabilization temperature for tDRH whilst halving the quadruplex-stabilizing affinity, compared to NSC273829. Overall, the molecules had selective affinity for the tDRH sequence compared to other forms of DNA, particularly the standard DNA duplex (cDD).

MD studies (**S4, ESI**) were undertaken using AMBER v11 to rationalise the stabilisation observed in the FRET melting assay. It was observed that compounds with an even number of methylene spacers (*i.e.*, compounds **2**, **4** and **6)** could *bis*-intercalate between the bases more effectively compared to compounds with an odd number of methylene spacers (*i.e.*, **1**, **3** and **5**), which supported the experimental observations (**Table 1**). More specifically, MD simulations suggested that in the case

of compounds with an even number of methylene spacers, both benzimidazole moieties are in the correct orientation to intercalate into the tDRH structure (**Figure 5a**), whereas an odd number of methylene spacers results in one benzimidazole moiety pointing out of the minor groove (**Figure 5b**). Among the molecules examined, **4** provided the best stabilisation of the tDRH duplex sequence compared to NSC273829 (**S4, ESI**). A 10 ns implicit solvent molecular dynamics simulation showed that the compound remained restrained over the sequence 5′-GTTAG-3′ for the duration of the simulation due to favourable van der Waal's interactions with the central 5′-TTA-3′ triplet (**Figure 5a**). Simulations of **6** illustrated that this molecule binds to tDRH in a comparable manner to **4**. A similar intercalation interaction does not occur with NSC273829, which instead sits in the minor groove (Figure **S4.2**, **ESI**). The greater stabilisation of tDRH by **4** and **6** compared to NSC273829 was supported by free energy of binding calculations (kcal/mol) of the intercalation of **Library 1** molecules and NSC273829 with tDRH and cDD sequences in implicit solvent (**Table 2**). **Library 1** molecules were also assessed against a control DNA sequence in which the uracil (**U**) of the tDRH duplex sequence was replaced with thymine (**T**), in which case none of the designed molecules provided any notable stabilisation in the FRET melting assay.

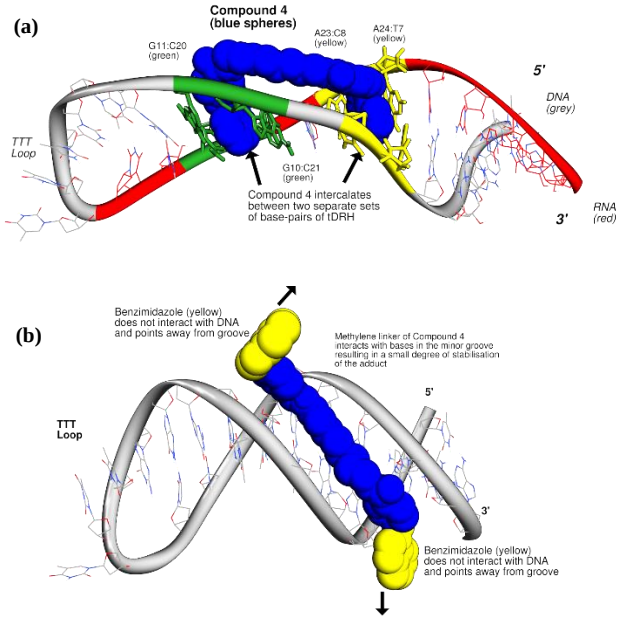

Figure 5: (a) Snapshot of a 10ns implicit solvent molecular dynamics simulation of **4** (blue spheres) interacting with tDRH. Both benzimidazole moieties intercalate into the sequence, one between G11:C20 and G10:C21 (green) and the second between A23:T8 and A24:T7 (yellow). **(b)** Snapshot of a 10 ns explicit solvent molecular dynamics simulation of **4** (blue and yellow spheres) interacting with the DNA duplex sequence (cDD). Both benzimidazole moieties orient away from the DNA groove.

The molecular modelling results also supported the data from the biophysical experiments in suggesting that these molecules are unlikely to *bis*-intercalate into the control duplex DNA in an effective manner. First, a significant difference in topologies exists between the minor grooves of the DNA and tDRH duplexes. The former has an enhanced curvature, whereas the tDRH minor groove is much flatter in its architecture. This difference in topology allows isohelical structures such as the polyamides (*e.g.*, distamycin) to bind with high affinity in the duplex DNA minor groove, whereas NSC273829 and related molecules do not possess the appropriate curvature to interact with duplex DNA. Second, **Library 1** molecules were designed to possess a chair-like shape to match the topology of tDRH, and should induce binding through a dual mechanism of action (*i.e.*, minor groove binding and intercalation). Therefore, the designed molecules may be able to selectively interact with the tDRH duplex through both modes.

Table 2: Free energy of binding calculations (kcal/mol).

Compound	Free Energy of Binding (MM-PBSA†)		
	tDRH	**cDD**	**RRH**
4	-51.49	-48.27	-36.49
6	-53.62	-49.44	-34.33
NSC273829	-29.76	-35.59	-32.76

†Molecular Mechanics Poisson Boltzmann Surface Area.

Molecular models and free energy of binding calculations suggest that the binding of **Library 1** molecules to RNA will most likely result in significant disorder of the nucleic acid structure, as the shape of compounds of this type is not consistent with RNA topology. In the case of compound **4**, both benzimidazole moieties point away from the RNA groove and do not participate in binding (Figure **S4.3, ESI**). This binding conformation results in base-pair displacement and strand separation which is reflected in the calculated free energies (*i.e.*, less favourable than those obtained for cDD and tDRH).

Next, a circular dichroism (CD) study was carried out to evaluate the interaction of the synthesized ligands with the tDRH structure (**S5, ESI**). Little CD information is available in the literature relating to DNA/RNA hybrid duplexes, especially those originating in the telomeric region of DNA. The CD spectrum of the tDRH duplex in Tris-HCl (50 mM) was initially measured, and an intense positive CD signal was observed at ~271 nm, along with a small negative CD signal at ~237 nm (Figure **S5.1, ESI**). The CD spectrum of the control DNA duplex (cDD) showed a strong positive CD signal at ~269 nm and a negative signal at ~241 nm (**Figure S5.1**, **ESI**). Compounds **4** and **6** induced significant changes in the positive CD signal at 237 nM, but did not cause any notable changes to the negative CD signal (**Figure 6**). A dose-dependent enhancement of the positive CD signal correlated well with the FRET melting results. For example, for **4** (**Figure 6a**), dose-dependent red shifts of between 1.2 to 4 nM were observed after addition of up to 5 equivalents of the ligand. Interestingly, hypochromic shifts were observed for **4** which is unusual, as red shifts are usually associated with hyperchromic shifts. An isoelliptic point for the spectrum was observed for the different concentrations of the compounds, suggesting that the molecules are working through a similar but specific mode of action (*i.e.*, *bis*-intercalation).

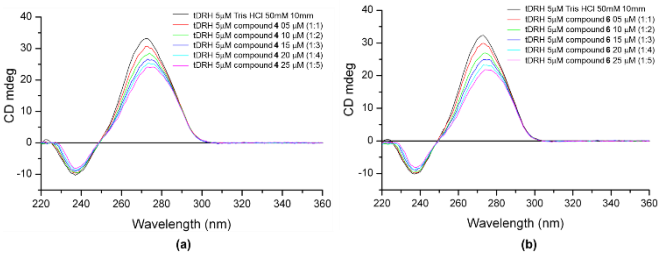

Figure 6: CD spectra of compound **4 (a)** and **6 (b)** with tDRH (5 μM) in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration.

Compound **6** produced a very similar CD titration profile (**Figure 6b**) with dose-dependent red shifts and hypochromic effects. Addition of up to 5 equivalents of **4** and **6** to the cDD sequence did not produce any change in the CD signal (**Figure S5.2**, **ESI**). The shifts and changes in intensity observed for the tDRH sequence and the lack of interaction observed for the cDD sequence supported both the FRET-melting data and the molecular modelling results.

The FRET melting assay, molecular modelling and dynamics studies along with the CD titration results support the hypothesis that **Library 1** molecules stabilise the tDRH structure by *bis*-intercalation. The **Library 1** molecules were also found to interact with telomeric quadruplex-forming DNA (*e.g.*, F21T) (**Figure 1c**) in the FRET studies (**Table 1**). Molecular modelling studies suggest that they may be able to bind on the periphery of tDNA quadruplex structures (**Figure S4.4**, **ESI**) due to their curvature. If developed as telomerase inhibitors, their telomeric quadruplex-binding properties may also contribute to their potency, as molecules that stabilise telomeric quadruplex structures are known to inhibit the telomerase enzyme.

To provide further evidence for the proposed mechanism of action of binding to the tDRH duplex, and to generate structure-activity relationship (SAR) data, molecular modelling was used to guide the design of a number of molecules lacking the key features of **Library 1** compounds (**ESI, Table S6.1**). Key features were considered to be: i) the flexibility afforded by the methylene spacer, ii) the single methylene spacers between the benzimidazole and phenylene moieties, and iii) the planar structure of the terminal benzimidazole moieties. Thus, in **Library 2** the methylene linkers were partly replaced with phenylene groups to provide rigidity to the linker, and in **Library 3** the flexible methylene linker of the molecules was retained, but the terminal single methylene groups were removed and the terminal benzimidazole moieties replaced with 6-fluoro-2-methylquinoline moieties. Finally, in **Library 4**, an additional cyclohexyl ring was introduced (*via* 7-fluoro-1,2,3,4-tetrahydroacridine) to remove planarity and prevent intercalation (**Figure 7**). For **Library 3** and **4** molecules, an identical length of central methylene spacers was used to produce the *bis*-amide products (**Table S6.2, S6.3 and S6.4, ESI**). The library members were synthesized using similar amide coupling procedures to those described earlier.

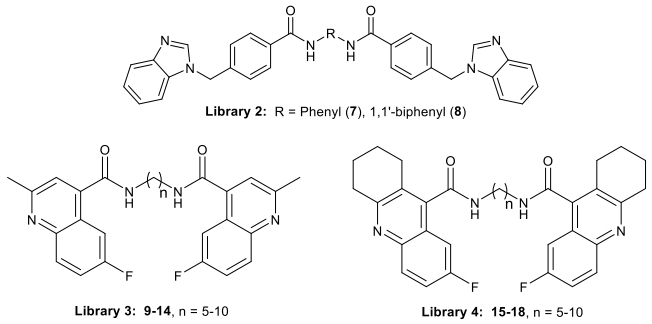

Figure 7: General structures of **Libraries 2**, **3** and **4**.

All molecules in **Libraries 2**, **3** and **4** were evaluated for their ability to stabilise tDRH based on FRET melting and CD. None produced any notable stabilisation (*i.e.*, ≥1°C) of the tDRH duplex at the highest concentration tested (*i.e.*, 5 μM). All molecules were found to be CD inactive. This confirmed the importance of the key features of **Library 1** molecules for interaction with the tDRH duplex. These results were consistent with the molecular modelling predictions, which suggested that the molecules lack the appropriate shape to intercalate into the tDRH structure, and should have only weak interactions in the minor groove of the A-form structure. The modelling also highlighted the importance of the flexibility of the central methylene spacers for *bis*-intercalation, as the rigid and planar phenylene (*i.e.*, **7**) and biphenylene (*i.e.*, **8**) containing compounds with identical terminal units to those used in **Library 1**, were unable to intercalate.

Finally, **Library 1** molecules were evaluated for their cytotoxicity in two tumour cell lines using a MTT assay (**S7, ESI**). In MDA-MB-231 and NCI H1975 cells (**Table 3**), **4** and **6** produced low micromolar IC_{50} values after 48 hours incubation. The potency of compounds with an even number of methylene spacers was marginally higher compared to those with an odd number, consistent with the results of the biophysical and computational experiments. However, further studies are required to demonstrate that the observed cytotoxicity is associated with telomere maintenance in tumour cells.

Table 3: IC_{50} values of **Library 1** molecules after 48 hours incubation

Compound	IC_{50} Value (μM) * Mean (n = 3)	
	MDA-MB-231	NCI H1975
2	3.4	> 100
3	4.2	16.8
4	1.1	11.8
5	9.2	15.3
6	4.2	1.5

In summary, a novel chemical scaffold has been identified capable of *bis*-intercalating into the tDRH structure, but with no affinity for duplex DNA due to its 3-dimensional shape. This offers an opportunity to selectively target the telomeric DNA/RNA duplex for therapeutic purposes while not affecting normal genomic DNA. The ligands described here can be used as a starting point to generate more potent molecules while potentially retaining selectivity. In particular, there is the possibility of creating diversity in the terminal heterocyclic groups, which may also be modified to optimise drug-like characteristics.

‡ **ESI:** Literature molecules (**S1**), Synthesis, Purification and Analysis (**S2**), FRET Assay (**S3**), Molecular modelling studies (**S4**), Modelling simulations of compound **6** (**S4.1**), NSC273829 (**S4.2**) and compound **3** on quadruplex structure (**S4.4**), CD data (**S5.2**, **S5.3**, **S5.4**), Synthesized molecules and their characterization (**S6**), Synthesised Library 2, 3 and 4 molecules (**S6.2**, **S6.3** and **S6.4**), Cytotoxicity assay (**S7**), ^{13}C-NMR and HRMS data (**S8**).

References

1. N. N. Shaw and D. P. Arya, *Biochimie*, 2008, **90**, 1026-1039.
2. C. L. Schildkraut, J. Marmur and P. Doty, *Journal of molecular biology*, 1961, **3**, 595-IN516.
3. A. Rich, *Proceedings of the National Academy of Sciences of the United States of America*, 1960, **46**, 1044-1053.

4. G. Milman, R. Langridge and M. J. Chamberlin, *Proceedings of the National Academy of Sciences of the United States of America*, 1967, **57**, 1804-1810.
5. G. Gavory, M. Farrow and S. Balasubramanian, *Nucleic Acids Res*, 2002, **30**, 4470-4480.
6. W. H. Gmeiner, W. Cui, D. E. Konerding, P. A. Keifer, S. K. Sharma, A. M. Soto, L. A. Marky and J. W. Lown, *J Biomol Struct Dyn*, 1999, **17**, 507-518.
7. R. Francis, C. West and S. H. Friedman, *Bioorg Chem*, 2001, **29**, 107-117.
8. C. West, R. Francis and S. H. Friedman, *Bioorg Med Chem Lett*, 2001, **11**, 2727-2730.
9. J. Ren and J. B. Chaires, *Biochemistry*, 1999, **38**, 16067-16075.
10. C. M. Barbieri, T. K. Li, S. Guo, G. Wang, A. J. Shallop, W. Pan, G. Yang, B. L. Gaffney, R. A. Jones and D. S. Pilch, *J Am Chem Soc*, 2003, **125**, 6469-6477.
11. S. Rangarajan and S. H. Friedman, *Bioorg Med Chem Lett*, 2007, **17**, 2267-2273.
12. N. N. Shaw, H. Xi and D. P. Arya, *Bioorg Med Chem Lett*, 2008, **18**, 4142-4145.
13. R. T. Wheelhouse, N. C. Garbett, N. J. Buurma and J. B. Chaires, *Angew Chem Int Edit*, 2010, **49**, 3207-3210.
14. K. M. Rahman, K. Tizkova, A. P. Reszka, S. Neidle and D. E. Thurston, *Bioorg Med Chem Lett*, 2012, **22**, 3006-3010.
15. B. K. Sinha, R. M. Philen, R. Sato and R. L. Cysyk, *J. Med. Chem.*, 1977, **20**, 1528-1531.
16. S. St George, J. V. Bishop, R. G. Titus and C. P. Selitrennikoff, *Antimicrob Agents Chemother*, 2006, **50**, 474-479.
17. T. M. Evans, J. M. Gardiner, N. Mahmood and M. Smis, *Bioorg Med Chem Lett*, 1997, **7**, 409-412.
18. S. H. Nile, B. Kumar and S. W. Park, *Chemical biology & drug design*, 2013, **82**, 290-295.
19. A. W. White, N. J. Curtin, B. W. Eastman, B. T. Golding, Z. Hostomsky, S. Kyle, J. Li, K. A. Maegley, D. J. Skalitzky, S. E. Webber, X. H. Yu and R. J. Griffin, *Bioorg Med Chem Lett*, 2004, **14**, 2433-2437.